Test Assignment :
Polynomial Regression

January 22, 2018

Objectives:

- In this assignment, you will study the polynomial regression problem. You are required to fit a polynomial model to some given data points with different orders by finding the analytical solution of a least square error problem.

General Notes:

- You should use the numpy and matplotlib libraries to complete this assignment. You are encouraged to look up the numpy and matplotlib documentation for useful utility functions to help you complete the programming portion of the assignment.
- Full points are given for complete solutions, including justifying the choices or assumptions you made to solve each question. Both your complete source code and written up report should be included in the final submission.
- Homework assignments are to be solved by yourself or in groups of two. You are encouraged to discuss the assignment with other students, but you must solve it within your own group. Make sure to be closely involved in all aspects of the assignment. If you work in a group, please indicate the contribution percentage from each group member at the beginning of your report.

Introduction

Given the data points $(x_1, t_1), \dots, (x_N, t_N)$, fitting these data points using a polynomial function of order M is given by

$$y(x, \mathbf{w}) = w_0 + w_1 x + w_2 x^2 + \cdots + w_M x^M = \sum_{m=0}^{M} w_m x^m$$

The polynomial function can be written in vectorized form as follows:

$$\mathbf{y}(\mathbf{x}, \mathbf{w}) = \mathbf{w}^T \boldsymbol{\phi}(\mathbf{x})$$

where $\mathbf{w} = [w_0, w_1, w_2, \ldots, w_M]^T \in \mathbf{R}^{M+1}$, $\mathbf{y} \in \mathbf{R}^N$, $\mathbf{x} \in \mathbf{R}^N$, and $\phi(\mathbf{x}) \in \mathbf{R}^{(M+1)\times N}$. $\boldsymbol{\phi}(\mathbf{x})$ is given by

$$\boldsymbol{\phi}(\mathbf{x}) = \begin{bmatrix} 1 & 1 & 1 & \ldots & 1 \\ x_1 & x_2 & x_3 & \ldots & x_N \\ x_1^2 & x_2^2 & x_3^2 & \ldots & x_N^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_1^M & x_2^M & x_3^M & \ldots & x_N^M \end{bmatrix}$$

The optimal weights $\hat{\mathbf{w}}$ are chosen by minimizing the square error between the prediction of the model $y(x_n, \mathbf{w})$ and the true target t_n, in other words,

$$\hat{\mathbf{w}} = \arg\min_{\mathbf{w}} \|\mathbf{t} - \mathbf{w}^T \boldsymbol{\phi}(\mathbf{x})\|_2^2 = \arg\min_{\mathbf{w}} \sum_{n=1}^{N} (t_n - \mathbf{w}^T \boldsymbol{\phi}(x_n))^2$$

The analytical solution to the previous minimization equation is given by

$$\hat{\mathbf{w}} = (\boldsymbol{\phi}(\mathbf{x})\boldsymbol{\phi}(\mathbf{x})^T)^{-1}\boldsymbol{\phi}(\mathbf{x})\mathbf{t}$$

1 Polynomial Regression

1. You are given a set of 10 data points as shown in the next Figure stored in *datapoints.npy* file. Make sure to place the file in the same directory of the program you will write. Use the following function to load the data points:

   ```
   def load_data():
       data = np.load('./datapoints.npy')
       train_data = data[:,0]
       train_target = data[:,1]
       return train_data, train_target
   ```

   You are required to write a python program to fit a polynomial function of order $M = \{1, 2, \ldots, 9\}$ to these data points by finding the optimal weights analytically. Report the optimum weights for $M = \{1, 2, \ldots, 9\}$.

2. use the function *np.linspace(-2, 2, 100)* as an input to your model and plot the prediction of your model along with the given data points. Use a single plot for each model order.

3. Based on your plots, what is the model order that best fits the data points (i.e. capture the structure of the data points) and the model order that overfits the data points? Comment on your choice.

4. **Regularization:**
For the over-fitting case, use a regularization term $\lambda = 0.1$ to find the optimum solution and report the optimum weights in this case. Repeat step 2 again and comment on your plots. Please note that the optimum weights in case of regularization is given by

$$\hat{\mathbf{w}} = (\boldsymbol{\phi}(\mathbf{x})\boldsymbol{\phi}(\mathbf{x})^T + \lambda\mathbf{I})^{-1}\boldsymbol{\phi}(\mathbf{x})\mathbf{t}$$

where $\mathbf{I}$ is the identity matrix $\in \mathbf{R}^{(M+1)\times(M+1)}$ with the first element is zero (because the bias term w_0 is not regularized), i.e.

$$\mathbf{I} = \begin{bmatrix} 0 & 0 & 0 & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{bmatrix}$$